Pattern and trends in occupational attainment in first jobs in the Netherlands, 1930-1995: OLS regression versus conditional multinomial logistic regression

Ganzeboom, H.B.G.

published in
Journal of the Royal Statistical Society. Series A. Statistics in Society
2003

DOI (link to publisher)
10.1111/j.1467-985X.2004.00336.x

document version
Publisher's PDF, also known as Version of record

J. R. Statist. Soc. A (2003)
166, *Part* 1, *pp.* 63–84

Patterns and trends in occupational attainment of first jobs in the Netherlands, 1930–1995: ordinary least squares regression *versus* conditional multinomial logistic regression

Jos A. G. Dessens, Wim Jansen, Harry B. G. Ganzeboom and
Peter G. M. van der Heijden
Utrecht University, the Netherlands

[Received April 2001. Revised April 2002]

Summary. This paper brings together the virtues of linear regression models for status attainment models formulated by second-generation social mobility researchers and the strengths of log-linear models formulated by third-generation researchers, into fourth-generation social mobility models, by using conditional multinomial logistic regression (CMLR). These CMLR models are capable of capturing the discrete and multidimensional nature of social mobility patterns (a characteristic of third-generation output) while reducing the number of parameters leading to parsimonious models (a characteristic of second-generation output). Using data from eight pooled surveys in the Netherlands, an extended Blau–Duncan status attainment model is formulated and analysed. The corresponding CMLR model is formulated incorporating general and specific inheritance effects. The final CMLR model gives a relatively parsimonious description of Dutch mobility patterns, similar to the extended Blau–Duncan model, at the same time offering the possibility of including specific effects where necessary. Effects of gender and education appear to be too complex to be captured by a single parameter.

Keywords: Conditional multinomial logistic regression; Log-linear models; the Netherlands; Social mobility; Status attainment

1. Class status and models: theoretical and methodological considerations

In the history of the analysis of social mobility, the pendulum has swung between continuous and discrete conceptions of social stratification (Ganzeboom *et al.*, 1991). The background of this dialectic has been varied, with theoretical, methodological and convenience arguments all pushing the balance. In the first generation of social mobility research, a continuous conception of social stratification prevailed. Researchers as different as Lloyd Warner (1949), Glass (1954) and Svalastoga (1959) assumed that social classes could best be considered as ‘social layers’ that were somehow ordered unidimensionally along a vertical axis, and movement between the layers was—almost implicitly—assumed to be governed solely by someone’s position in this hierarchy. Although questions were occasionally raised about the specific nature, location, number and boundaries of the various social groups, the agenda of the first generation is best characterized by one of its main objectives: the construction of a valid and detailed continuous measure of occupational prestige.

Address for correspondence: Jos A. G. Dessens, Department of Methodology and Statistics, Interuniversity Center for Social Science Theory and Methodology, Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, the Netherlands.
E-mail: J.Dessens@fss.uu.nl

 0964–1998/03/166063

Curiously, researchers of the first generation frequently used inherently discrete methods in analysing inflow and outflow of occupational categories in the context of social mobility tables. Most importantly, the comparative projects of that generation (Lipset and Zetterberg, 1959; Miller, 1960) used the non-manual–manual–farm trichotomy to achieve comparability between measures. As a conception of social stratification, this trichotomy approaches occupational prestige scales in a very vague way. Moreover, while analysing their tables, mobility researchers of the first generation found that some characteristics of their mobility tables were difficult to reconcile with a one-dimensional representation of statistical distributions, such as a correlation coefficient or some other measure of association. At best, such measures will average the pattern of association in a mobility table, but they cannot represent it fully. In particular, the overrepresentation of socially immobile individuals on the diagonal of a social mobility table and the differentiation of immobility between social classes (being particularly large among farmers and other self-employed groups) could not be modelled by the techniques of the 1950s.

One major contribution of the second generation of social mobility research, initiated by Blau and Duncan's (1967) status attainment model, was the use of path analysis models for the association between more than two variables. Most importantly, indirect effect calculations made it possible to quantify the role of education in social mobility and to disclose its reproductive and mobility promoting facets. This was achieved at the cost of assuming that all variables in the model are continuous. Later researchers, using the technique of linear structural equations, implicitly assume that the world satisfies a multivariate normal distribution that can be adequately summarized by using means, standard deviations and correlations. The immense advantage of this idea is that all relationships can be expressed in a few easily interpreted model parameters. However, a major disadvantage of second-generation models is that, by concentrating on only a few parameters to model the patterns of association, these estimated effects are highly sensitive to contextual influences. For this reason, from the very beginning, status attainment models were criticized on the basis of their overly simple conception of social stratification as a single and continuous hierarchy. Indeed, some researchers in this tradition suggested that (continuous) status measures of stratification were inferior to (discrete) class measures in their associations with dependent variables, such as income or class consciousness (Kelley, 1973; Robinson and Kelley, 1979). Whereas discrete versions of variables were easily incorporated on the predictor side by using dummy variables, there were no simple ways of achieving this on the dependent side of the model, and it required a process of methodological developments to accomplish this.

The third generation of social mobility research (Hauser, 1978; Goldthorpe *et al.*, 1980) returned to tabular analysis, but this time using log-linear models as a modelling tool. This generation, and in particular the comparative work, culminating in the 'Comparative analysis of social mobility in industrial nations' (CASMIN) project (Erikson and Goldthorpe, 1992) conclusively confirmed that social mobility patterns cannot be summarized adequately by a single parameter. Using the tool of log-linear models to separate marginal and association effects, it was shown that the failure to do so in second-generation models could (and sometimes did) lead to erroneous conclusions. Moreover, the patterns of association in mobility tables were found to be highly complex with distinct components that could vary independently with exogenous conditions. For instance, the CASMIN core model, espoused by Erikson and Goldthorpe, uses eight parameters to model 36-dimensional association in a 7×7 contingency table. Comparative versions of the model are highly complicated by the fact that the core model cannot be upheld without structural adaptations.

Although admittedly these third-generation models have conclusively established the discrete and multidimensional nature of patterns of social mobility, the scope and lucidity of the approach are open to criticism, which may explain why the popularity of these models has dwindled in the recent literature. One distinct problem of log-linear analyses in the manner of the CASMIN project is that they yield so many parameters that they make conclusions hard to digest and interpret. Another related criticism is that models which consume so many degrees of freedom lack statistical power, which in turn may explain why so few comparative conclusions have been established. Finally, the most important drawback of the third-generation models is that they have effectively brought us back to the study of bivariate associations, at the same time neglecting the multivariate structure of social mobility. In other words: log-linear analysis has brought us more detailed insights about increasingly less.

The problem of coefficients that are difficult to interpret and the lack of statistical power in third-generation models has been countered by the use of scaled association models, such as Goodman's row-and-column effect models RCII (Goodman, 1979) and Xie's uniform difference models (Xie, 1992). Although these adaptations go a long way towards simplifying models and increasing statistical power, they cannot overcome the problem that log-linear models for the relationship between origin class and destination class are intrinsically bivariate.

Research in social mobility and status attainment in the Netherlands has closely followed the international trends in this field. The Dutch were among the first generation of mobility researchers, with contributions by Van Heek *et al.* (1958) and van Tulder (1962). They produced a prestige scale and social mobility tables according to international guidelines, only to find their data discarded in the international comparative studies because they did not fit into the non-manual–manual–farm framework. The second generation included few Dutch sociologists, but there was a marked reawakening at the dawn of the log-linear models era. Ganzeboom and De Graaf (1984) and Ganzeboom *et al.* (1987) constructed mobility tables that covered the 1954–1977 and 1970–1985 periods respectively. Ignoring the then-current many parameter CASMIN model, they employed variations of the multiplicative RCII model for the association (Goodman, 1979) to model trends in social mobility. Although methodologically unexceptional, their results tended to deviate from results reported for other countries. In particular, they found that intergenerational occupational mobility in the Netherlands tends to go up over time, which conflicts with conclusions elsewhere, where the constant social fluidity and constant flux were found to alternate (Erikson and Goldthorpe, 1992). That is, researchers concluded either that there was little variation in intergenerational mobility patterns, or that variations displayed no meaningful trends.

Fourth-generation models of social mobility should improve on earlier generations by combining the best of all worlds: they should be multivariate, allowing for conclusions to be drawn on both indirect effects and discrete categories, while using a limited number of degrees of freedom to be sensitive to contextual conditions. Multinomial logistic regression models that are suitable for this purpose were first introduced about 20 years ago (Logan, 1983), but their practical application in mobility research has been rather limited until now.

The aim of this paper is to show how the second-generation Blau–Duncan status attainment model and the third-generation log-linear models can be transformed into a fourth-generation multinomial logistic regression model that allows for testing similar hypotheses for both Blau–Duncan and log-linear models. In the following section, the Blau–Duncan model will be extended and some of its shortcomings explained. In Section 3, a description of the Dutch data to be used in the analyses is given. Sections 4 and 5 contain the results of the extended Blau–Duncan model and various log-linear models. After introducing the conditional multinomial logistic regression (CMLR) model, results from the CMLR analyses are given in Sections 7 and 8. The final section presents our conclusions.

2. Extending the Blau–Duncan model and linking second- and third-generation mobility models

De Graaf and Luijkx (1993) extended the Blau–Duncan occupational attainment model for the Netherlands. Of particular interest here is that, besides father's occupational attainment and respondent's education, they added respondent's gender, and his or her labour market experience. They assumed that the effect of a respondent's education on his or her occupational attainment differs between men and women. They also assumed that the effect of father's occupational status on a respondent's occupational status differs between men and women. Furthermore, the effect of a respondent's education on his or her occupational status was supposed to vary with the year of entry into the labour market. Here, it is expected that the effect of education should increase over time as a result of educational expansion. Next, it is suggested that the effect of father's occupational status on a respondent's occupational status varies with the year of entry into the labour market as well as with educational level. The effect of father's occupational status will diminish over time, and the effect of father's occupational status will be less for people with longer education than for people having short educational careers. This is in line with Mare's findings, that the effect of social background will decrease with the length of the educational career (Mare, 1980; Rijken, 1999).

Although the extended Blau–Duncan model described above may help us to understand the status attainment process, some of the shortcomings of this model remain. First, occupational status attainment is based on an *a priori* ranking of the occupational categories, which does not come from the data itself and may not adequately represent it. Second, it is impossible to incorporate inheritance and/or social immobility effects. Specifically, if we want to incorporate an overall immobility effect and/or specific inheritance effects for small proprietors and self-employed farmers in the extended Blau–Duncan model, there is no obvious way to do so. Third, as noted in the previous section, the path model coefficients are assumed to be homogeneous for each of the occupational categories, a hypothesis that cannot be adequately tested.

Switching to a CMLR model avoids these shortcomings. As mentioned in the preceding section, in the third generation of social mobility research, the continuous occupational status attainment variables were replaced by categorical versions, and regression techniques by log-linear analysis. Highly successful models were developed within this generation of mobility studies, such as quasi-independence models and (scaled) multiplicative models of association (RCII models) between the categorical origin and destination variables. The effects of immobility and inheritance, and scale multiplicative association (Goodman, 1979) are easily incorporated into any CMLR model. At the same time, the model allows for continuous covariates, a major advantage of the Blau–Duncan approach. In the CMLR model, a separate effect for each category of the dependent variable (categorical destination class in this case) must be estimated for each of the predictor variables. Instead of having a single parameter for the effect of educational level, there will be as many parameters for the effect of education as there are categories of destination.

Typical applications of CMLR analysis may suffer from the same problems as early log-linear models for tabular data: they come with an unwieldy number of parameters that are both difficult to interpret and difficult to remember. It is only recently that researchers in social mobility have learned ways to reduce the number of parameters in these models to a manageable and informative set (Breen, 1994; DiPrete, 1990; Hendrickx and Ganzeboom, 1998) by modelling the association between origin and destination in terms of the predictor variables.

In the analysis reported in this paper, a detailed view is given of trends in effects on occupational status attainment associated with the first jobs of women and men in the Netherlands over a considerable period of time. After a description of the data, the analysis begins from the second-generation-extended Blau–Duncan model, passing through the third-generation log-linear analysis and ending in the corresponding fourth-generation CMLR model. Our ultimate goal is to achieve a balance between providing the amount of detail that is needed to describe the data adequately and maximum parsimony. Integrating immobility, inheritance, uniform association and RCII models within CMLR will allow more precise and detailed answers to the original hypotheses and extensions, derived from the Blau–Duncan model.

3. Description of the data

The data are taken from eight nationwide surveys, conducted in the Netherlands between 1982 and 1996. The surveys are part of the *International Stratification and Mobility File* (ISMF) collection. This is a collection of standardized sample surveys with detailed information on occupational titles of the respondents, and of their fathers, educational levels and various background characteristics. Because of this standardization, data from the ISMF are highly appropriate for the various analyses to be presented in the next sections. Descriptions of the surveys used here can be found in Table 1. The ISMF is maintained by H. B. G. Ganzeboom and D. J. Treiman (`http://www.fss.uu.nl/soc/hg/ismf/index.htm`). Although all the surveys are of sufficient quality to sustain detailed analysis, they vary somewhat in the exact definition of the sample and the measurement of the variables. The analyses here are restricted to men and women in the 21–64 years age range, on the assumption that the vast majority of these respondents will have had the opportunity to enter the labour market. We made no accommodation for the fact that, in each of the surveys, those with the highest educational levels in the youngest generation are under-represented, because they are still in the educational system at 21 years of age.

Using the procedures, set out by Ganzeboom and Treiman (1996), occupational information was reduced to an eight-category version of the Erikson–Goldthorpe–Portocarero (EGP) typology of social class (Erikson *et al.*, 1979). The EGP class scheme is as follows: 1, large proprietors, higher professionals and managers; 2, lower professionals and managers; 3, routine non-manual workers; 4, small proprietors with and without employees; 5, lower grade technicians and manual supervisors, and skilled manual workers; 6, unskilled and semi-skilled workers; 7, agricultural workers; 8, self-employed farmers.

The surveys varied in the measurement of education, a variable that we shall assume to be continuous throughout the analysis. Education is converted into a comparable format by matching the respective categories in each of the surveys with the minimally required number of years to complete a specific educational level. Beside respondent's EGP code and education, and father's EGP code, we shall use the information on respondent's age and gender.

For the status attainment parts of the analysis, the EGP categories have been recoded into the international socioeconomic index (ISEI) (Ganzeboom *et al.*, 1992; Ganzeboom and Treiman, 1996). Each EGP category has been scaled according to the mean ISEI value (1, >68.6; 2, >58.5; 3, >47.0; 4, >37.4; 5, >35.8; 6, >28.8; 7, >16.7; 8, >26.7). The correlation between the scaled EGP and ISEI is 0.885, between the scaled and unscaled EGP 0.978 and between the ISEI and the unscaled EGP 0.866. From each respondent's age and educational career in years (EDYR), his or her labour market entry year (EYR) is derived. EYR has been centred at 1965. The relevant summary statistics for the variables used in the analyses are given in Table 2.

Table 1. Description of the data used in the analyses

Producer	*Year*	*Title*	*Distributor*
Werkgroep Nationaal Kiezersonderzoek	1982	Nationaal Kiezersonderzoek (Dutch Parliamentary Election Panel Study)	Steinmetz Archive, Amsterdam
M. I. L. Gijsberts and H. B. G. Ganzeboom	1996	Sociale Ongelijkheid in Nederland 1996 (Social inequality in the Netherlands 1996)	Steinmetz Archive, Amsterdam
International Social Justice Project	1991	Perceptions of justice, international merged data set	
International Social Survey Programme Zentralarchiv für Empirische Sozialforschung, University of Cologne	1986, 1987	Social inequality	Interuniversity Consortium for Political and Social Research, Ann Arbor
W. C. Ultee and H. Sixma	1982	National Prestige Survey	Steinmetz Archive, Amsterdam
W. C. Ultee and H. B. G. Ganzeboom	1992, 1993	Familie-enquete Nederlandse Bevolking (Family survey of the Dutch population)	Vakgroep Sociologie KUN
J. Weesie and H. B. G. Ganzeboom	1994	Huishoudens in Nederland (Households in the Netherlands) (telephone survey)	
K. Wittebrood and M. J. Tervoert Nederlands Studiecentrum voor Criminaliteit en Rechtshandhaving, Leiden	1996	Netherlands survey on criminality and law enforcement	

4. Results from the second-generation extended Blau–Duncan model

Results of a regression analysis of respondent's occupational status (ISEI-scaled EGP) as the dependent variable on variables related to his or her social background and father's characteristics as main effects are shown in Table 3 (model 1). In model 2, interactions are added between education (EDYR) and gender (FEMALE), between father's occupational status (FISEI) and gender, between education and entry year (EYR), between father's occupational status and entry year, and between father's occupational status and education.

The interaction terms EDYR ∗ FEMALE and FISEI ∗ FEMALE test the hypotheses that the effects of education and of father's occupational prestige differ between men and women. The interaction terms EDYR ∗ EYR and FISEI ∗ EYR embody hypotheses about trends in the effects of father's occupational status and respondent's education on his or her occupational status over time. It is expected that the effect of father's occupational status on respondent's occupational status should diminish over time, and that the effect of education on occupational status should increase. The interaction term FISEI ∗ EDYR refers to Mare's findings

Table 2. Summary statistics for the variables in the analyses

Variable	*N*	*Minimum*	*Maximum*	*Mean*	*Standard deviation*
Respondent's EGP	7621	1	8	3.9	1.7
Respondent's ISEI	7621	13	88	42.9	14.0
Respondent's EGP (scaled)	7621	16.7	68.6	42.9	12.4
Respondent's father's FEGP	7621	1	8	4.3	2.1
Respondent's father's FISEI	7621	13	88	42.3	15.6
Respondent's education in years	7621	0	21	11.0	3.1
Respondent's age	7621	21	64	40.4	11.4
Respondent's labour market entry year (centred at 1965)	7621	−35	33	1.4	13.9

that the effect of father's occupational status decreases with the length of the educational career.

The results in Table 3 indicate that the effect of education on occupational status is stronger for men than for women. The effect of father's occupational status on respondent's occupational status does not differ between men and women, or with increasing duration of educational career. The influence of father's occupational status on occupational status of the respondent diminishes over time, whereas the effect of father's occupational status does not diminish with the length of the educational career as was predicted by Mare.

As noted in Section 2, several shortcomings of this extended Blau–Duncan model can be formulated:

(a) occupational prestige is ascertained by an *a priori* ranking of the occupational categories and is not based on the data;
(b) it is not possible to incorporate inheritance effects, e.g. an overall inheritance effect and specific inheritance effects for small proprietors and self-employed farmers;

Table 3. Results from the extended Blau–Duncan model for the Netherlands†

	Model 1, B (t-value)	*Model 2, B (t-value)*
Main effects		
Father's occupational status (FISEI)	0.16 (19.4)	0.16 (4.9)
Education in years (EDYR)	1.67 (39.9)	1.69 (12.7)
FEMALE	2.69 (10.9)	7.78 (7.8)
Labour market entry year (EYR)	−0.003 (−0.3)	0.08 (2.2)
Interactions		
FISEI ∗ FEMALE		−0.03 (−1.8)
FISEI ∗ EDYR		0.002 (0.7)
FISEI ∗ EYR		−0.003 (−4.3)
EDYR ∗ FEMALE		−0.34 (−4.0)
EDYR ∗ EYR		0.004 (1.3)
R^2	0.298	0.303

†The dependent variable is the respondent's scaled EGP category. Model 1 is without and model 2 is with interactions.

(c) path model coefficients are assumed to be homogeneous for each of the occupational categories, and there is no clear way to test this assumption.

A way to overcome these shortcomings is to return to the categorical representation of occupational status, as proposed by Erikson *et al.* (1979). Social mobility within their framework can be linked to changes in market and work situations, rather than between status levels. It seems obvious to adopt this approach. Cross-classifying the origin and destination occupational categories leads to a contingency table. Log-linear models are highly suitable for modelling the association between origin and destination. In the next section we shall show the strength of these log-linear models, while demonstrating the impossibility of incorporating other predictor variables from the extended Blau–Duncan model.

5. Results from third-generation log-linear models

In log-linear models, the cell frequencies f_{ij} of contingency tables are modelled. In social mobility research, the two-dimensional contingency table of origin (indexed by i) and destination (indexed by j) class categories is typically the research object. Table 4 present two cross-classifications for the Netherlands.

Table 4. Cross-classification of origin and destination (first occupation) for the Netherlands, 1982–1996†

FEGP	*Results for the following EGPs:*								*Total*
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	
All respondents									
1	95	277	273	12	70	85	8	1	821
2	89	368	447	18	122	151	9	3	1207
3	45	163	327	12	90	135	7	2	781
4	19	116	264	40	129	149	10	1	728
5	39	177	517	16	371	393	22	1	1536
6	29	149	468	22	294	584	33	6	1585
7	2	28	51	5	39	98	31	3	257
8	17	94	177	13	95	162	73	75	706
Total	335	1372	2524	138	1210	1757	193	92	7621
Male respondents									
1	65	126	94	8	56	59	7	1	416
2	70	200	173	13	104	84	7	3	654
3	34	92	161	7	82	82	6	2	466
4	16	60	133	34	117	92	9	1	462
5	34	90	239	8	340	238	19	0	968
6	22	64	177	13	274	337	29	4	920
7	2	7	28	4	39	57	30	3	170
8	14	50	72	4	86	93	63	72	454
Total	257	689	1077	91	1098	1042	170	86	4510

†1, large proprietors, higher professionals and managers; 2, lower professionals and managers; 3, routine non-manual workers; 4, small proprietors with and without employees; 5, lower grade technicians, manual supervisors and skilled manual workers; 6, unskilled and semi-skilled workers; 7, agricultural workers; 8, self-employed farmers.

Testing for association in such a cross-classification starts from the model of independence

$$\log(m_{ij}) = u + u_{1(i)} + u_{2(j)}, \tag{1}$$

with m_{ij} the expected cell frequency under the model, u a constant, $u_{1(i)}$ the main effect of origin (FEGP) and $u_{2(j)}$ the main effect of destination (EGP). In social mobility research, the independence model is bound to be rejected. There is a consistent relationship between origin and destination.

Instead of modelling the association by using the saturated model

$$\log(m_{ij}) = u + u_{1(i)} + u_{2(j)} + u_{12(ij)}, \tag{2}$$

more parsimonious models can be used to capture the multidimensional association structure in a testable way. A uniform association parameter could be estimated, and/or various 'immobility' and 'inheritance' effects. The number of different models is restricted only by the number of degrees of freedom between the model of independence and the saturated model. An extensive overview of log-linear models for mobility tables is given in Hout (1983). As is well known (Heath, 1981), the social mobility process differs for men and women. This leads to differences in uniform association and/or inheritance parameters for men and women. The simple log-linear models must be extended with the gender variable and applied to the three-dimensional contingency table of origin class (FEGP) by destination class (EGP) by gender (FEMALE) for the Netherlands. Results for the various models are given in Table 5.

The final two models do not scale the EGP and FEGP categories in a uniform but in an optimal way. Model 14 allows for a different scaling of origin and destination categories, whereas model 15 constrains them to be equal. The model of scaled uniform association with equality

Table 5. Log-linear analysis results for various (im)mobility and inheritance models for the Netherlands†

Model	*Deviance*	*Degrees of freedom*	*BIC*‡
1, EGP + FEGP + FEMALE	2735	112	1734
2, (EGP + FEGP) ∗ FEMALE	1556	98	680
3, model 2 + GENDIAG	1008	97	141
4, (EGP + FEGP + GENDIAG) ∗ FEMALE	993	96	135
5, model 2 + IDIAG	765	90	−39
6, (EGP + FEGP + IDIAG) ∗ FEMALE	736	82	3
7, model 2 + UNIFASS	607	97	−260
8, (EGP + FEGP + UNIFASS) ∗ FEMALE	598	96	−260
9, model 4 + UNIFASS	435	95	−414
10, (EGP + FEGP + GENDIAG + UNIFASS) ∗ FEMALE	434	94	−406
11, (EGP + FEGP + IDIAG + UNIFASS) ∗ FEMALE	295	80	−420
12, model 10 + FARM + SPROP	329	92	−493
13, (EGP + FEGP + GENDIAG + UNIFASS + FARM + SPROP) ∗ FEMALE	322	90	−482
14, (EGP + FEGP + GENDIAG + RCII) ∗ FEMALE + FARM + SPROP	242	80	−473
15, (EGP + FEGP + GENDIAG + HOMRCII) ∗ FEMALE + FARM + SPROP	254	86	−515

†EGP, origin category; FEGP, destination category; GENDIAG, overall diagonal parameter; IDIAG, specific parameter for each diagonal cell; UNIFASS, uniform association; SPROP, inheritance parameter for small proprietors; FARM, inheritance parameter for self-employed farmers; RCII, scaled uniform association; HOMRCII, RCII with equal scaling.
‡The BIC statistic is defined as [deviance − $\ln(N)\times$ degrees of freedom].

Table 6. Scaling parameters for models 14 and 15

Category	*FEGP*	*EGP*	*HOMRCII*
1	−0.57	−0.53	−0.57
2	−0.39	−0.39	−0.40
3	−0.24	−0.15	−0.17
4	−0.03	−0.08	−0.05
5	0.11	−0.07	0.09
6	0.25	0.17	0.22
7	0.51	0.67	0.58
8	0.36	0.23	0.31

constraints is clearly superior to the non-scaled uniform association model 12. The estimated, normalized, scaling parameters for models 14 and 15 are given in Table 6.

Model 15 states that the overall immobility and the scaled uniform association pattern differ for men and women, that inheritance processes are present, but that the specific inheritance parameters for self-employed farmers and small shopkeepers do not depend on gender. These results are known to be dependent on variables like education, and year of entry into the labour market. If we want to include these variables, it is necessary to categorize these continuous variables into a relatively small number of categories. Otherwise the multidimensional table will quickly become too sparse. CMLR models allow for the incorporation of continuous variables as covariates. At the same time, all the well-known log-linear model features are preserved in these CMLR models.

6. *Intermezzo*: multinomial and conditional multinomial logistic regression models

The first log-linear model to be considered within the framework of CMLR models is that of statistical independence between father's and respondent's occupational category: model 1. The corresponding multinomial logistic regression model with the first destination category as the reference category is

$$\log(m_{ij}/m_{i1}) = u_{2(j)} - u_{2(1)} = u_{2(j)} - 0 = \alpha_j. \tag{3}$$

This model may be seen as consisting of seven ($j = 2, 3, \ldots, 8$) simultaneous logit equations with the design or model matrix having identical blocks of non-zero elements. In this case the columns α_1–α_8 of the design or model matrix (given in Table 7) refer to a model with a constant only.

Fitting this multinomial logistic regression model by using any statistical software package that uses the multinomial logit algorithm will result in parameter estimates for the α-parameters, which are identical with the $u_{2(j)}$-parameters from the log-linear model of independence, given that the same reference category is used. As was clear from the previous section, the independence model will not fit. One reason for this is that this model ignores inheritance processes. Most prominently these inheritance processes play a role in the case of the small proprietor and self-employed farming origins. Intergenerational transfer of capital goods most probably takes place within these categories. Therefore the obvious log-linear model to be considered next is a model with inheritance parameters for small proprietor and self-employed farming origins:

Table 7. Model matrix for the CMLR models†

	α_2	α_3	α_4	α_5	α_6	α_7	α_8	$d_4 * SPROP$	$d_8 * FARM$	ν
$\log(m_{12}/m_{11})$	1	0	0	0	0	0	0	0	0	$\kappa_1(\lambda_2-\lambda_1)$
$\log(m_{22}/m_{21})$	1	0	0	0	0	0	0	0	0	$\kappa_2(\lambda_2-\lambda_1)$
$\log(m_{32}/m_{31})$	1	0	0	0	0	0	0	0	0	$\kappa_3(\lambda_2-\lambda_1)$
$\log(m_{42}/m_{41})$	1	0	0	0	0	0	0	0	0	$\kappa_4(\lambda_2-\lambda_1)$
$\log(m_{52}/m_{51})$	1	0	0	0	0	0	0	0	0	$\kappa_5(\lambda_2-\lambda_1)$
$\log(m_{62}/m_{61})$	1	0	0	0	0	0	0	0	0	$\kappa_6(\lambda_2-\lambda_1)$
$\log(m_{72}/m_{71})$	1	0	0	0	0	0	0	0	0	$\kappa_7(\lambda_2-\lambda_1)$
$\log(m_{82}/m_{81})$	1	0	0	0	0	0	0	0	0	$\kappa_8(\lambda_2-\lambda_1)$
$\log(m_{13}/m_{11})$	0	1	0	0	0	0	0	0	0	$\kappa_1(\lambda_3-\lambda_1)$
$\log(m_{23}/m_{21})$	0	1	0	0	0	0	0	0	0	$\kappa_2(\lambda_3-\lambda_1)$
$\log(m_{33}/m_{31})$	0	1	0	0	0	0	0	0	0	$\kappa_3(\lambda_3-\lambda_1)$
$\log(m_{43}/m_{41})$	0	1	0	0	0	0	0	0	0	$\kappa_4(\lambda_3-\lambda_1)$
$\log(m_{53}/m_{51})$	0	1	0	0	0	0	0	0	0	$\kappa_5(\lambda_3-\lambda_1)$
$\log(m_{63}/m_{61})$	0	1	0	0	0	0	0	0	0	$\kappa_6(\lambda_3-\lambda_1)$
$\log(m_{73}/m_{71})$	0	1	0	0	0	0	0	0	0	$\kappa_7(\lambda_3-\lambda_1)$
$\log(m_{83}/m_{81})$	0	1	0	0	0	0	0	0	0	$\kappa_8(\lambda_3-\lambda_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\log(m_{44}/m_{41})$	0	0	1	0	0	0	0	1	0	$\kappa_4(\lambda_4-\lambda_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\log(m_{18}/m_{11})$	0	0	0	0	0	0	1	0	0	$\kappa_1(\lambda_8-\lambda_1)$
$\log(m_{28}/m_{21})$	0	0	0	0	0	0	1	0	0	$\kappa_2(\lambda_8-\lambda_1)$
$\log(m_{38}/m_{31})$	0	0	0	0	0	0	1	0	0	$\kappa_3(\lambda_8-\lambda_1)$
$\log(m_{48}/m_{41})$	0	0	0	0	0	0	1	0	0	$\kappa_4(\lambda_8-\lambda_1)$
$\log(m_{58}/m_{51})$	0	0	0	0	0	0	1	0	0	$\kappa_5(\lambda_8-\lambda_1)$
$\log(m_{68}/m_{61})$	0	0	0	0	0	0	1	0	0	$\kappa_6(\lambda_8-\lambda_1)$
$\log(m_{78}/m_{71})$	0	0	0	0	0	0	1	0	0	$\kappa_7(\lambda_8-\lambda_1)$
$\log(m_{88}/m_{81})$	0	0	0	0	0	0	1	0	1	$\kappa_8(\lambda_8-\lambda_1)$

†Independence model (α_1–α_8); inheritance model for small proprietors and self-employed farmers (α_1–α_8 and d_4 * SPROP and d_8 * FARM); scaled uniform association model (α_1–α_8 and ν).

$$\log(m_{ij}) = u + u_{1(i)} + u_{2(j)} + d_4 * \text{SPROP} + d_8 * \text{FARM} \tag{4}$$

with $d_4 = 1$ if $i = j = 4$, and $d_4 = 0$ otherwise, and $d_8 = 1$ if $i = j = 8$, and $d_8 = 0$ otherwise.

The corresponding model in terms of the logits with $j = 1$ as the reference category is

$$\begin{aligned} \log(m_{ij}/m_{i1}) &= u_{2(j)} - u_{2(1)} + d_4 * \text{SPROP} + d_8 * \text{FARM} \\ &= u_{2(j)} + d_4 * \text{SPROP} + d_8 * \text{FARM} \\ &= \alpha_j + d_4 * \text{SPROP} + d_8 * \text{FARM}. \end{aligned} \tag{5}$$

The design matrix for this model corresponds to columns α_1–α_8 and d_4 * SPROP and d_8 * FARM in Table 7. Now, the blocks in the design or model matrix, corresponding to the various destinations, are no longer identical, as can be noted from the columns that go with the inheritance parameters d_4 and d_8. Reaching the self-employed farming destination is conditional on coming from a self-employed farming origin. Therefore, this logit model of inheritance is not a multinomial logistic regression model but a CMLR model.

This 'inheritance' model does not give a satisfactory fit either. Clearly, further modelling of the association between origin and destination is necessary. A scaled uniform association parameter ν is added to the model. Goodman's model RCII is fitted, in which the κ_i (origin scale values) and λ_j (destination scale values) must also be estimated. Goodman's model RCII is a log-multiplicative model and requires the alternating fixing of the κ_i and λ_j estimates in fitting the model:

$$\log(m_{ij}) = u + u_{1(i)} + u_{2(j)} + d_4 * \text{SPROP} + d_8 * \text{FARM} + \kappa_i * \lambda_j * \nu \tag{6}$$

with $d_4 = 1$ if $i = j = 4$, and $d_4 = 0$ otherwise, and $d_8 = 1$ if $i = j = 8$, and $d_8 = 0$ otherwise.

The corresponding conditional multinomial logistic regression model looks like

$$\begin{aligned} \log(m_{ij}/m_{i1}) &= u_{2(j)} - u_{2(1)} + d_4 * \text{SPROP} + d_8 * \text{FARM} + \kappa_i(\lambda_j - \lambda_1)\nu \\ &= u_{2(j)} + d_4 * \text{SPROP} + d_8 * \text{FARM} + \kappa_i(\lambda_j - \lambda_1)\nu \\ &= \alpha_j + d_4 * \text{SPROP} + d_8 * \text{FARM} + \kappa_i(\lambda_j - \lambda_1)\nu. \end{aligned} \tag{7}$$

Adding a uniform association parameter means adding a column to the model matrix in Table 7, which equals the $\kappa_i(\lambda_j - \lambda_1)$-terms.

In fitting a CMLR model, the model matrix must be expanded, because (some of) the entries of the model matrix are dependent on a specific destination category. Each observation is expanded into as many observations as there are categories of the dependent variable. In the social mobility example the dependent variable is the eight-category destination class. To demonstrate this expansion we show the first three lines of the data matrix before (Table 8, part (a)) and after (Table 8, part(b)) expansion.

The respondent's occupational category (EGP = 7 for the first observation) in the expanded model matrix can be obtained from the value of ID in the row for which EGP = 1. As was noted earlier, the entries in the column in the design matrix corresponding to the scaled uniform association parameter (model RCII) can be calculated as $\kappa_i(\lambda_j - \lambda_1)$. The general immobility parameter is referred to as GENDIAG.

It is immediately clear from the expansion that individual characteristics such as education and gender are of no use in estimating the probabilities for each of the categories of the dependent variable, because these variables do not vary between the records for a specific observation after expansion. Only variables that vary with respect to the categories of the dependent variable, e.g. the attractiveness of each category to the individual, can be used in estimating the CMLR model. Individual characteristics such as education and gender can be incorporated in CMLR models by allowing for the possibility that the effect of these variables may differ for each of the categories of the dependent variable. This can be accomplished by forming interactions between these individual variables and dummy variables that represent the constant in the logit equations.

7. Modelling the association within the conditional multinomial logistic regression context

For comparing the results from the status attainment model and the CMLR model, it is necessary to have a categorical variable that is substantively equivalent to the ISEI-scaled EGP variable. An obvious candidate for such a categorical variable is the eight-category EGP variable.

Next, we must identify the comparable effects in the status attainment and the CMLR models. In the CMLR model, the (log-) odds of attaining occupational category j relative to a reference occupational category is predicted. In contrast with the FEMALE effect in the status attainment model, the main effect of, for example, FEMALE in the CMLR analysis is represented by seven parameters (i.e. the number of destination categories minus 1). The CMLR model includes the

Table 8. Part of the data matrix

FEGP	*EGP*	*GENDIAG*	*RCII*	*EDYR*	*FEMALE*	*ID*	*FEGP*	*EGP*	*GENDIAG*	*RCII*	*EDYR*	*FEMALE*
(a) Before expansion						*(b) After expansion*						
6	7	0	0.16	4	1	1	6	0	0	0	4	1
						2	6	0	0	0.10	4	1
						3	6	0	0	0.07	4	1
						4	6	0	0	0.11	4	1
						5	6	0	0	0.08	4	1
						6	6	0	1	0.16	4	1
						7	6	1	0	0.16	4	1
						8	6	0	0	0.14	4	1
4	1	0	0	4	0	1	4	1	0	0	4	0
						2	4	0	0	0.24	4	0
						3	4	0	0	0.17	4	0
						4	4	0	1	0.27	4	0
						5	4	0	0	0.20	4	0
						6	4	0	0	0.39	4	0
						7	4	0	0	0.41	4	0
						8	4	0	0	0.36	4	0
3	3	1	−0.03	5	1	1	3	0	0	0	5	1
						2	3	0	0	−0.04	5	1
						3	3	1	1	−0.03	5	1
						4	3	0	0	−0.05	5	1
						5	3	0	0	−0.04	5	1
						6	3	0	0	−0.07	5	1
						7	3	0	0	−0.07	5	1
						8	3	0	0	−0.07	5	1

main effects for the predictor variables FEMALE, EDYR and EYR, and the interaction effects EDYR $*$ FEMALE and EDYR $*$ EYR.

The main and interaction effects of FISEI in the status attainment model must be represented differently here. The effect of FISEI on scaled EGP in the status attainment framework is concerned with the association between these variables. Within the CMLR context, the effect of FEGP on EGP must be estimated by the association between these variables. This can be done by fitting Goodman's model RCII, containing a parameter for the association between the scaled occupational categories of origin and destination, next to the main and some of the interaction effects parameters, corresponding to those from the status attainment model. The interaction terms in the status attainment model which include FISEI can be estimated in the CMLR context by modelling the scaled association parameter in terms of the variables involved in the FISEI interaction terms. For example, the interaction effect FISEI $*$ FEMALE in the status attainment model corresponds to the interaction between FEMALE and the scaled association parameter within the CMLR context.

Within the CMLR framework, the RCII model is as follows (the parameters σ and ϕ parallel the parameters κ and λ, introduced earlier; they may differ in the identification restrictions):

$$\log(m_{ij}/m_{i1}) = \alpha_j - \alpha_1 + \sum_{k=1}^{K} \beta_{jk} X_k + \nu\sigma_i(\phi_j - \phi_1) \tag{8}$$

with restrictions

$$\sum \phi_j = \sum \sigma_i = 0$$

and

$$\sum \phi_j^2 = \sum \sigma_i^2 = 1.$$

In contrast with the regression analysis in Section 4, it is now possible to model immobility and inheritance processes. Here, three more parameters are added to the model: a general immobility parameter, a specific inheritance parameter for children of small proprietors and a specific inheritance parameter for children of self-employed farmers.

The scale values estimated from the log-linear homogeneous RCII model 15 in Section 5 are used in the next modelling step. We calculate a scaled uniform association vector $U = \phi * \sigma$, and add this vector to the CMLR model matrix. Moreover, interactions of this association parameter U with the predictors FEMALE, EDYR and EYR are added to the model. Results are given in Table 9. Model A is the counterpart of the original extended Blau–Duncan status attainment model. In model B, immobility and inheritance effects, and trends in these effects, are added.

As noted earlier, for each of the main and some of the interaction effects there are seven parameters, representing the effects of the corresponding predictor variables on the log-odds of reaching destination category ($j \neq 1$) over the reference category (1). In this analysis, the first EGP category (large proprietors, higher professionals and managers) has been chosen as the reference category.

In all the models, there is a large uniform association effect, meaning that there is a substantial relationship between origin and destination categories. This finding is in line with the FISEI effect in the status attainment model. The interaction effects U $*$ FEMALE, U $*$ EDYR and U $*$ EYR show the same pattern as in the status attainment model; the uniform association between origin and destination categories diminishes over time. As to the main effect of FEMALE, it can be seen that, in contrast with the regression analysis, there are now four significant

Table 9. Results of CMLR analyses for the Netherlands; without (model A) and with (model B) immobility and inheritance parameter estimates; and patterns and trends in immobility and inheritance parameters (model C)

	Parameters (z-values) for the following models:		
	Model A	*Model B*	*Model C*
Effects father's occupation			
Uniform association (*U*)	5.56 (6.7)	4.35 (5.3)	4.37 (5.3)
U ∗ FEMALE	0.13 (0.3)	0.41 (1.0)	0.37 (0.9)
U ∗ EDYR	−0.15 (−2.1)	−0.13 (−1.9)	−0.13 (−1.9)
U ∗ EYR	−0.07 (−4.5)	−0.06 (−4.1)	−0.06 (−3.6)
Immobility and inheritance effects			
General immobility parameter (GENDIAG)		0.32 (7.9)	0.32 (7.8)
Inheritance small proprietors (SPROP)		0.86 (4.3)	0.68 (2.8)
Inheritance self-employed farmers (FARM)		3.04 (11.0)	2.92 (10.1)
Patterns and trends in immobility and inheritance effects			
GENDIAG * FEMALE		−0.16 (−2.3)	−0.15 (−2.2)
GENDIAG * EYR			−0.001 (−0.3)
SPROP * EYR			−0.03 (−1.6)
FARM * EYR			−0.04 (−1.9)
Main and interaction effects			
EDYR			
Occupational category 2 *versus* occupational category 1	−0.18 (−6.1)	−0.18 (−6.1)	−0.18 (−6.1)
Occupational category 3 *versus* occupational category 1	−0.45 (−15.0)	−0.45 (−15.1)	−0.45 (−15.1)
Occupational category 4 *versus* occupational category 1	−0.55 (−10.7)	−0.56 (−10.8)	−0.56 (−10.7)
Occupational category 5 *versus* occupational category 1	−0.62 (−19.5)	−0.62 (−19.4)	−0.62 (−19.4)
Occupational category 6 *versus* occupational category 1	−0.63 (−19.8)	−0.63 (−19.7)	−0.63 (−19.7)
Occupational category 7 *versus* occupational category 1	−0.63 (−13.7)	−0.64 (−13.8)	−0.64 (−13.8)
Occupational category 8 *versus* occupational category 1	−0.55 (−9.9)	−0.59 (−9.7)	−0.60 (−9.8)
FEMALE			
Occupational category 2 *versus* occupational category 1	3.49 (3.7)	3.69 (3.8)	3.69 (3.8)
Occupational category 3 *versus* occupational category 1	4.28 (4.5)	4.47 (4.6)	4.47 (4.6)
Occupational category 4 *versus* occupational category 1	3.60 (3.0)	3.92 (3.2)	3.94 (3.2)
Occupational category 5 *versus* occupational category 1	1.42 (1.4)	1.67 (1.6)	1.66 (1.6)
Occupational category 6 *versus* occupational category 1	3.33 (3.5)	3.57 (3.6)	3.57 (3.6)

(*continued*)

Table 9 (*continued*)

	Parameters (z-values) for the following models:		
	Model A	*Model B*	*Model C*
Occupational category 7 *versus* occupational category 1	1.58 (1.2)	1.77 (1.4)	1.77 (1.4)
Occupational category 8 *versus* occupational category 1	1.98 (1.1)	3.30 (1.7)	3.38 (1.7)
EYR			
Occupational category 2 *versus* occupational category 1	0.06 (2.2)	0.05 (2.0)	0.05 (2.0)
Occupational category 3 *versus* occupational category 1	0.04 (1.5)	0.03 (1.3)	0.03 (1.3)
Occupational category 4 *versus* occupational category 1	0.03 (0.8)	0.03 (0.8)	0.04 (1.0)
Occupational category 5 *versus* occupational category 1	0.05 (2.0)	0.05 (1.8)	0.05 (1.8)
Occupational category 6 *versus* occupational category 1	−0.005 (−0.2)	−0.01 (−0.4)	−0.01 (−0.4)
Occupational category 7 *versus* occupational category 1	−0.02 (−0.5)	−0.02 (−0.6)	−0.02 (−0.6)
Occupational category 8 *versus* occupational category 1	0.01 (0.2)	0.01 (0.2)	0.05 (1.1)
EDYR ∗ FEMALE			
Occupational category 2 *versus* occupational category 1	−0.16 (−2.6)	−0.17 (−2.8)	−0.17 (−2.8)
Occupational category 3 *versus* occupational category 1	−0.21 (−3.3)	−0.22 (−3.4)	−0.22 (−3.4)
Occupational category 4 *versus* occupational category 1	−0.23 (−2.4)	−0.24 (−2.5)	−0.25 (−2.5)
Occupational category 5 *versus* occupational category 1	−0.18 (−2.4)	−0.19 (−2.5)	−0.19 (−2.5)
Occupational category 6 *versus* occupational category 1	−0.18 (−2.7)	−0.19 (−2.8)	−0.19 (−2.8)
Occupational category 7 *versus* occupational category 1	−0.15 (−1.5)	−0.17 (−1.6)	−0.17 (−1.6)
Occupational category 8 *versus* occupational category 1	−0.23 (−1.4)	−0.37 (−1.9)	−0.38 (−1.9)
EDYR ∗ EYR			
Occupational category 2 *versus* occupational category 1	−0.004 (−2.0)	−0.003 (−1.8)	−0.003 (−1.8)
Occupational category 3 *versus* occupational category 1	−0.003 (−1.5)	−0.002 (−1.3)	−0.002 (−1.3)
Occupational category 4 *versus* occupational category 1	−0.004 (−1.3)	−0.003 (−1.1)	−0.004 (−1.2)
Occupational category 5 *versus* occupational category 1	−0.004 (−1.9)	−0.003 (−1.8)	−0.003 (−1.8)
Occupational category 6 *versus* occupational category 1	0.002 (1.0)	0.002 (1.2)	0.002 (1.2)
Occupational category 7 *versus* occupational category 1	0.003 (1.2)	0.003 (1.2)	0.003 (1.2)
Occupational category 8 *versus* occupational category 1	−0.002 (−0.6)	−0.000 (−0.1)	−0.002 (−0.4)
Constant			
Occupational category 2 *versus* occupational category 1	3.63 (8.7)	—	—
Occupational category 3 *versus* occupational category 1	7.50 (17.9)	—	—

(*continued*)

Table 9 (*continued*)

	Parameters (z-values) for the following models:		
	Model A	*Model B*	*Model C*
Occupational category 4 *versus* occupational category 1	6.12 (9.9)	—	—
Occupational category 5 *versus* occupational category 1	9.33 (21.5)	—	—
Occupational category 6 *versus* occupational category 1	9.21 (21.3)	—	—
Occupational category 7 *versus* occupational category 1	7.15 (13.1)	—	—
Occupational category 8 *versus* occupational category 1	5.89 (9.0)	—	—
−2 log-likelihood	10258 (46 degrees of freedom)	10569 (50 degrees of freedom)	10576 (53 degrees of freedom)

positive effects (lower professionals and managers, routine non-manual, small proprietors and unskilled and semi-skilled manual), and three non-significant: skilled manual, farm labourers and self-employed farmers. This result is far more detailed than the corresponding FEMALE parameter in the regression model. The main effects of EDYR are much more closely in line with the EDYR effect on scaled EGP in the regression model in Section 4. However, from occupational category 3 to occupational category 8, the effects are almost identical. In model D, this result is used in forming separate contrasts for occupational category 2 and for occupational categories 3–8. The parameter estimates for these specific contrasts are as follows: occupational category 2 *versus* occupational category 1, −0.17 (−6.1); occupational categories 3–8 *versus* occupational category 1, −0.54 (−18.7); −2 log-likelihood (48 degrees of freedom), 10432.

The main non-significant effects of EYR are in line with the non-significant effect of this variable in the regression analysis. The interaction effects EDYR * FEMALE and EDYR * EYR also resemble their counterparts in the regression analysis. There is a general immobility effect, and specific inheritance effects for small proprietors and self-employed farmers. No trend in immobility is found; nor are there any trends in the specific inheritance processes. A striking feature of Table 9 is that the CMLR analyses lead to many parameters for main and interaction effects. The question may be raised whether this number of parameters can be further reduced. Stereotyped ordered regression (SOR) models will perform this task.

8. Conditional multinomial logistic regression model with immobility and inheritance parameters, stereotyped ordered regression restrictions and RCII scaling as an alternative for the extended Blau–Duncan model

In fitting a CMLR model, the effect of each covariate is represented by as many parameters as the number of categories of the dependent variable, say τ_1–τ_J. Of course, one of these parameters is redundant. Now, suppose that we have another covariate with parameters ω_1–ω_J. If we could assume that these two sets of parameters differ by only a constant k, i.e. $\tau_i = k\omega_i$, a substantial reduction in the number of parameters is obtained. This property may easily be extended to other covariates, each with their own constant. The idea of restricting the number

Table 10. Results of CMLR analyses with SOR restrictions on EDYR, EYR, EDYR * FEMALE and EDYR * EYR for the Netherlands†

	Parameters (z-values) for the following models:		
	Model A	*Model B*	*Model C*
Effects father's occupation			
U	4.48 (4.3)	4.55 (4.3)	4.47 (4.0)
U * FEMALE	0.34 (0.7)	0.32 (0.7)	
Occupational category 2 *versus* occupational category 1			−1.98 (−1.4)
Occupational category 3 *versus* occupational category 1			−3.46 (−1.0)
Occupational category 4 *versus* occupational category 1			−23.56 (−1.5)
Occupational category 5 *versus* occupational category 1			−3.36 (−0.5)
Occupational category 6 *versus* occupational category 1			4.72 (1.7)
Occupational category 7 *versus* occupational category 1			1.20 (0.6)
Occupational category 8 *versus* occupational category 1			2.09 (0.4)
U * EDYR	−0.14 (−1.6)	−0.15 (−1.6)	−0.14 (−1.4)
U * EYR	−0.07 (−3.4)	−0.07 (−3.1)	−0.07 (−3.0)
Immobility and inheritance effects			
General immobility parameter (GENDIAG)	0.33 (6.5)	0.32 (6.5)	0.32 (6.4)
Inheritance small proprietors (SPROP)	0.93 (4.1)	0.70 (2.9)	0.71 (2.9)
Inheritance self-employed farmers (FARM)	3.08 (9.9)	2.95 (9.5)	2.96 (9.4)
Patterns and trends in immobility and inheritance effects			
GENDIAG * FEMALE	−0.17 (−1.7)	−0.17 (−1.6)	−0.17 (−1.7)
GENDIAG * EYR		0.000 (−1.0)	0.000 (0.09)
SPROP * EYR		−0.03 (−2.1)	−0.03 (−2.1)
FARM * EYR		−0.02 (−1.5)	−0.02 (−1.4)
Main effect FEMALE			
Occupational category 2 *versus* occupational category 1	1.49 (3.7)	1.48 (3.7)	1.68 (3.8)
Occupational category 3 *versus* occupational category 1	2.17 (3.1)	2.15 (3.1)	2.33 (3.2)
Occupational category 4 *versus* occupational category 1	1.48 (1.7)	1.46 (1.8)	1.65 (1.9)
Occupational category 5 *versus* occupational category 1	−0.18 (−0.2)	−0.20 (−0.3)	−0.02 (−0.03)
Occupational category 6 *versus* occupational category 1	1.75 (2.1)	1.73 (2.1)	1.91 (2.3)
Occupational category 7 *versus* occupational category 1	0.16 (0.2)	0.14 (0.2)	0.37 (0.4)
Occupational category 8 *versus* occupational category 1	−0.40 (−0.1)	−0.40 (−0.1)	−0.16 (0.05)

(*continued*)

Table 10 (*continued*)

	Parameters (z-values) for the following models:		
	Model A	*Model B*	*Model C*
Main and interaction effects (SOR)			
EYR	−0.04 (−1.9)	−0.04 (−1.8)	−0.04 (−1.8)
EDYR	−0.67 (−19.8)	−0.66 (−19.8)	−0.66 (−19.8)
EDYR ∗ FEMALE	−0.08 (−1.3)	−0.07 (−1.2)	−0.08 (−1.3)
EDYR ∗ EYR	0.003 (1.6)	0.003 (1.6)	0.003 (1.5)
−2 log-likelihood	10497 (26 degrees of freedom)	10505 (29 degrees of freedom)	10524 (35 degrees of freedom)

†Parameter estimates were obtained by an alternating fitting algorithm. Therefore, the standard errors were estimated via a bootstrap procedure (100 replications).

of parameters in this way is known as SOR, developed by Anderson (1984). To illustrate the idea on our data, the main effects of EDYR and EYR and the interaction effects EDYR ∗ FEMALE and EDYR ∗ EYR are estimated by using SOR. Given restrictions $\Sigma\tau_j = 0$ and $\Sigma\tau_j^2 = 1$, the strength of the effect of variables like EDYR can be expressed in a single parameter β_{EDYR} instead of seven parameters as in the CMLR model (Hendrickx and Ganzeboom (1998), page 393). The use of the SOR model here could only be justified on the assumption that the 'causal mechanisms' of the various covariates on the destination class (EGP) are essentially comparable. The logit equation for the combined SOR and RCII model could be written as

$$\log(m_{ij}/m_{i1}) = \alpha_j - \alpha_1 + (\tau_j - \tau_1)\sum_{k=1}^{K}\beta_k X_k + \nu\sigma_i(\phi_j - \phi_1) \tag{9}$$

with restrictions

$$\sum \tau_j = \sum \phi_j = \sum \sigma_i = 0$$

and

$$\sum \tau_j^2 = \sum \phi_j^2 = \sum \sigma_i^2 = 1.$$

The parameter τ_j is a (normalized) scaling parameter for each of the occupational destination categories, common for all covariates X_k that have been included in the model.

As may be noted from the previous analyses, it is not advisable to represent the variable FEMALE by a single parameter. For the results reported in Table 10, seven contrasts for the variable FEMALE (models A and B) and another seven contrasts for the interaction between FEMALE and the uniform association parameter (model C) have been used.

Comparing models A, B and C, model A clearly stands out as the best model. It is far more parsimonious than the CMLR model without the SOR restrictions and can easily be compared with the results from the extended Blau–Duncan model. The effect of social origin on destination is captured by the significant association parameter (U), which states that there is an overall association between father's occupational category and respondent's occupational category. In addition, the immobility and inheritance parameters are substantial. Inheritance is highest among self-employed farmers. As in the CMLR model, it was also tested whether the immobility and inheritance effects became less important over time (model B). The interaction of these variables with EYR did show a significant negative effect for the small proprietors only. Inheritance seems to diminish over time within the small proprietors category, but not within the self-employed farming category. The gender variable FEMALE must be represented by seven parameters instead of one. Significant effects for this variable were found for the lower controllers, the non-manual, the small proprietors and the unskilled manual. Women have a higher chance of finishing in each of these occupational categories than men. No significant effects show up for the three occupational categories skilled manual, farm labourers and self-employed farmers.

In contrast with the findings from the Blau–Duncan status attainment model, there is no effect of gender on the association between father's occupation and respondent's occupation. EYR shows a negative effect on the association between father's occupation and respondent's occupation, a finding which parallels that from the Blau–Duncan model. The association between father's occupational category and respondent's occupational category diminishes over time. There is no interaction between EDYR and FEMALE, and also no significant interaction between EDYR and EYR. The first result conflicts with findings from the Blau–Duncan model. Lastly, the length of the educational career (EDYR) shows a consistent negative (but not significant) effect on the association between father's occupation and respondent's occupation. The direction of this effect implies a weak support of Mare's hypothesis.

9. Conclusion

CMLR models give a detailed description of the status attainment process when SOR restrictions, (scaled) association parameters, immobility and inheritance parameters are successively included in the model. The final model can be as detailed as required, and the details can reveal the effects of certain predictor variables on each of the logits formed by the categories of the dependent variable. In the case of the status attainment process for the Netherlands, gender effects are needed for each of the logits formed by the respondent's occupational category. Also, the effect of educational career seems to be most pronounced in the higher status categories, and less so in the other EGP categories. Other predictor variables and possible interactions, for which such a detailed description does not appear to be necessary, can be restricted to a single parameter using SOR restrictions. The association between father's occupational category and respondent's occupational category can be captured in a (scaled) association parameter, which turned out to be substantial for the Dutch data, and this parameter can be further modelled in terms of the predictor variables. Immobility and inheritance parameters may be added to the model. For the Dutch case, a general immobility parameter and two specific inheritance parameters, one for the small proprietors, and the other for the self-employed farmers, were needed. The strength of the second-generation status attainment models (few parameters) and the strength of the third-generation log-linear models (the inherently discrete and multidimensional nature of social mobility patterns) have thus been successfully combined in fourth-generation CMLR models.

Acknowledgements

We are grateful to Ruud Luijkx for checking some of our log-linear analysis results, and to the participants of the Utrecht stratification seminar for their comments on a draft version.

Appendix A: Software

In fitting CMLR models we used the Stata modules `mlogit` and `clogit`. For estimates of the standard errors of the SOR parameters, the Stata bootstrap procedure `bs` was used. For the Stata module `clogit`, Hendrickx (2000) wrote a set of macros to fit multiplicative models. These macros are available on his Web site `http://baserv.uci.kun.nl/~johnh/mcl/stata/`.

Stata 6.0 (StataCorp., 1999) was run on a personal computer under Windows 98.

Macros for fitting multinomial logistic regression models in GLIM are provided by Aitkin and Francis (1992); Kühnel (1990) described a maximum likelihood algorithm in SPSSx-Matrix. In the SPSS general log-linear module it is possible to incorporate cell covariates. In this approach the dimensions of the contingency table do not change; the means of the covariates for each cell in the contingency table are calculated and used as a vector in the design or model matrix. This is certainly not what most researchers aim at when incorporating covariates. For this and related problems we refer to Dessens *et al.* (1998). In the 9.0 and higher versions of SPSS, multinomial logistic regression models can be fitted by using the NOMREG module (SPSS, 1999). Here, covariates are correctly handled. The LIMDEP program (Greene, 1995) also provides CMLR modelling procedures, which were used by Breen (1994).

GLIM macros to fit multiplicative models within the log-linear framework can be found in Dessens *et al.* (1985).

References

Aitkin, M. and Francis, B. (1992) Fitting the multinomial logit model with continuous covariates in GLIM. *Comput. Statist. Data Anal.*, **14**, 89–98.

Anderson, J. A. (1984) Regression and ordered categorical variables (with discussion). *J. R. Statist. Soc.* B, **46**, 1–30.

Blau, P. M. and Duncan, O. D. (1967) *The American Occupational Structure*. New York: Free Press.

Breen, R. (1994) Individual level models for mobility tables and other cross-classifications. *Sociol. Meth. Res.*, **33**, 147–173.

De Graaf, P. M. and Luijkx, R. (1993) Trends in status attainment in the Netherlands from ascription to achievement. In *Solidarity of Generations: Demographic, Economic, and Social Change and Its Consequences* (eds H. A. Becker and P. L. J. Hermkens), pp. 437–466. Amsterdam: Thesis.

Dessens, J., Jansen, W. and van der Heijden, P. G. M. (1998) Using covariates in loglinear analysis with sampling zeros: a cautionary note. *Comput. Statist. Data Anal.*, **27**, 239–245.

Dessens, J., Jansen W. and Luijkx, R. (1985) Fitting log-multiplicative association models with GLIM. *GLIM Newslett.*, **11**, 28–34.

DiPrete, T. A. (1990) Adding covariates to loglinear models for the study of social mobility. *Am. Sociol. Rev.*, **55**, 757–773.

Erikson, R. and Goldthorpe, J. H. (1992) *The Constant Flux: a Study of Class Mobility in Industrial Societies*. Oxford: Clarendon.

Erikson, R., Goldthorpe, J. H. and Portocarero, L. (1979) Intergenerational class mobility in three Western European societies: England, France and Sweden. *Br. J. Sociol.*, **30**, 415–439.

Ganzeboom, H. B. G. and De Graaf, P. M. (1984) Intergenerational occupational mobility in the Netherlands in 1954 and 1977: a loglinear analysis. In *Social Stratification and Mobility in the Netherlands. A Collection of Recent Contributions to the Study of Social Inequality in a Modern Western Society: Data, Trends and Comparisons* (eds B. F. M. Bakker, J. Dronkers and H. B. G. Ganzeboom), pp. 71–90. Amsterdam: Stichting Interuniversitair Instituut voor Sociaal-wetenschappelijk Onderzoek.

Ganzeboom, H. B. G., De Graaf, P. M. and Treiman, D. J. (1992) A standard international socio-economic index of occupational status. *Socl Sci. Res.*, **21**, 1–56.

Ganzeboom, H., Luijkx, R., Dessens, J., De Graaf, P., De Graaf, N. D., Jansen, W. and Ultee, W. (1987) Intergenerationele klassenmobiliteit in Nederland tussen 1970 en 1985. *Mens Maatschap.*, **62**, 17–43.

Ganzeboom, H. B. G. and Treiman, D. J. (1996) Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. *Socl Sci. Res.*, **25**, 201–239.

Ganzeboom, H. B. G., Treiman, D. J. and Ultee, W. C. (1991) Comparative intergenerational stratification research: three generations and beyond. In *Annual Review of Sociology* (eds W. R. Scott and J. Blake), pp. 277–302. Palo Alto: Annual Reviews.

Glass, D. V. (ed.) (1954) *Social Mobility in Britain*. London: Routledge and Kegan Paul.

Goldthorpe, J. H., Llewellyn, C. and Payne, C. (1980) *Social Mobility and Class Structure in Modern Britain*. Oxford: Clarendon.

Goodman, L. A. (1979) Multiplicative models for the analysis of occupational mobility tables and other kinds of cross-classification tables. *Am. J. Sociol.*, **84**, 804–819.

Greene, W. (1995) *LIMDEP 7.0: User's Manual*. Bellport: Econometric Software.

Hauser, R. M. (1978) A structural model of the mobility table. *Socl Forces*, **56**, 919–953.

Heath, A. (1981) *Social Mobility*. Glasgow: Fontana.

Hendrickx, J. (2000) Special restrictions in multinomial logistic regression. *Stata Tech. Bull.*, **56**, 18–26.

Hendrickx, J. and Ganzeboom, H. B. G. (1998) Occupational status attainment in the Netherlands, 1920-1990: a multinomial logistic analysis. *Eur. Sociol. Rev.*, **14**, 387–403.
Hout, M. (1983) *Mobility Tables*. Beverly Hills: Sage.
Kelley, J. (1973) Causal chain models for the socioeconomic career. *Am. Sociol. Rev.*, **38**, 481–493.
Kühnel, S. M. (1990) Lassen sich mit SPSSx-Matrix anwenderspezifische Analyseprobleme lösen?: ein Anwendungstest am Beispiel der multinomialen logistischen Regression. *ZA-Inform.*, **27**, 89–109.
Lipset, S. M. and Zetterberg, H. L. (1959) Social mobility in industrial societies. In *Social Mobility in Industrial Society* (eds S. M. Lipset and R. Bendix), pp. 11–75. Berkeley: University of California Press.
Lloyd Warner, W., Meeker, M. and Eells, K. (1949) *Social Class in America*. Chicago: Science Research Associates.
Logan, J. A. (1983) A multivariate model for mobility tables. *Am. J. Sociol.*, **89**, 324–349.
Mare, R. D. (1980) Social background and school continuation decisions. *J. Am. Statist. Ass.*, **75**, 295–305.
Miller, S. M. (1960) Comparative social mobility: a trend report. *Curr. Sociol.*, **9**, 1–89.
Rijken, S. R. H. (1999) *Educational Expansion and Status Attainment: a Cross-National and Over-time Comparison*. Amsterdam: Thela Thesis.
Robinson, R. V. and Kelley, J. (1979) Class as conceived by Marx and Dahrendorf: effects on income inequality, class consciousness, and class conflict in the United States and Great Britain. *Am. Sociol. Rev.*, **44**, 38–58.
SPSS (1999) *SPSS Regression Models 9.0*. Chicago: SPSS.
StataCorp. (1999) *Stata Statistical Software: Release 6.0*. College Station: Stata Corporation.
Svalastoga, K. (1959) *Prestige, Class, and Mobility*. Copenhagen: Gyldendal.
van Tulder, J. J. M. (1962) *De Beroepsmobiliteit in Nederland van 1919 tot 1954: Een Sociaalstatistische Studie*. Leiden: Stenfert Kroese.
Van Heek, F., Vercruijsse, E. V. W., in't Veld-Langeveld, H. M., Kuiper, G., van Braam, A. and Korstanje, B. (1958) *Sociale Stijging en Daling in Nederland*. Leiden: Stenfert Kroese.
Xie, Y. (1992) The log-multiplicative layer effect for comparing mobility tables. *Am. Sociol. Rev.*, **57**, 380–395.